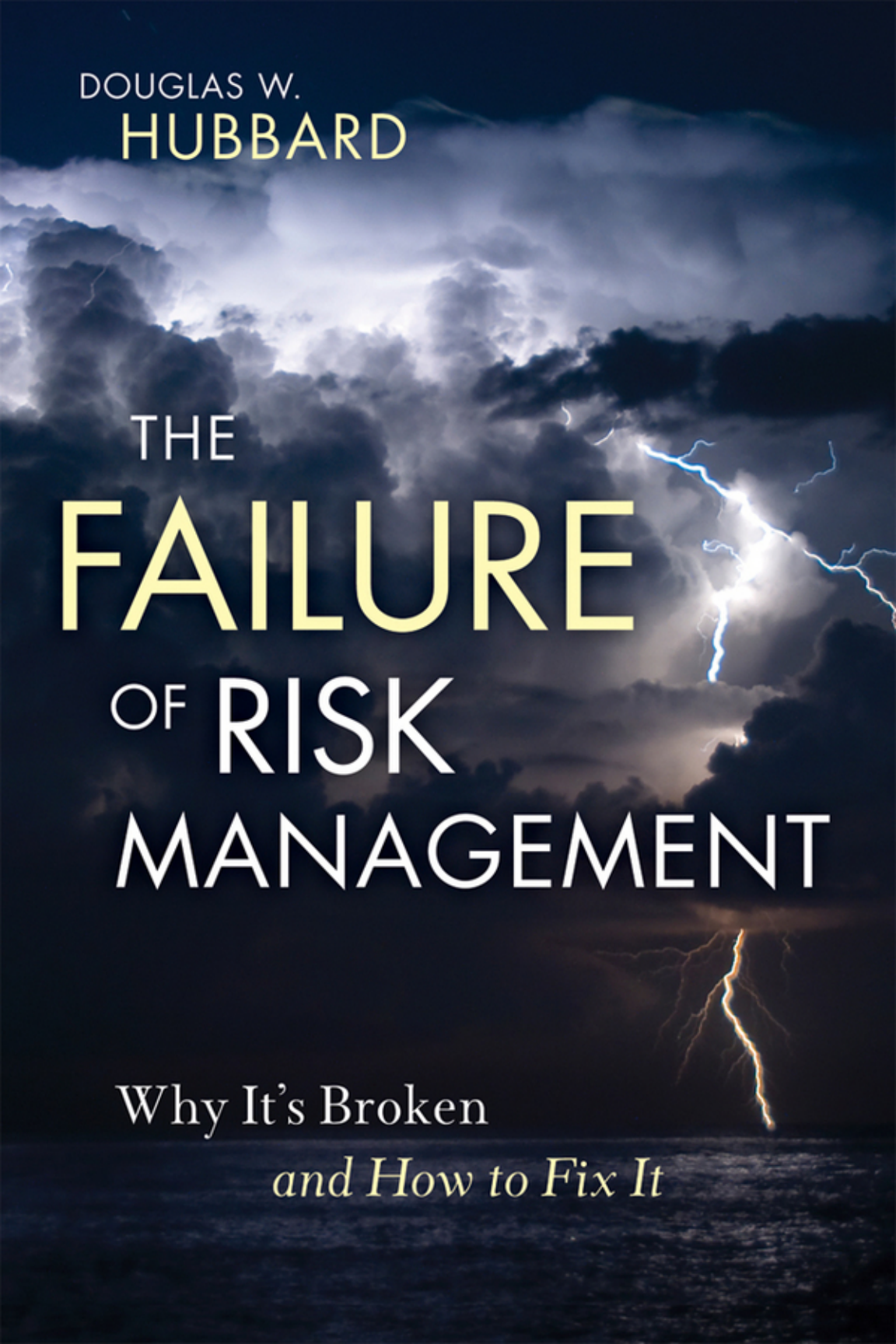
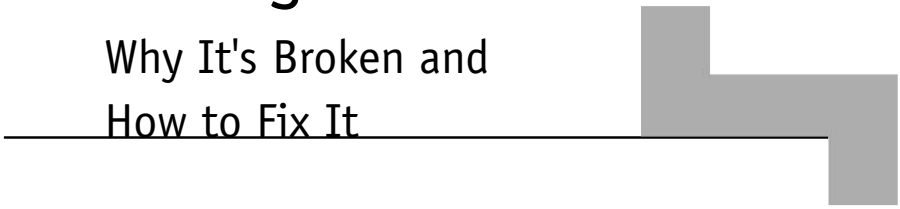

DOUGLAS W.
HUBBARD
THE
FAILURE
OF RISK
MANAGEMENT
Why It's Broken
and How to Fix It

The Failure of Risk Management

The Failure of Risk Management:

Why It's Broken and How to Fix It



Douglas W. Hubbard

Wiley

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data

Hubbard, Douglas W., 1962-
The failure of risk management : why it's broken and how to fix it / Douglas W. Hubbard.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-38795-5 (cloth)
1. Risk management. I. Title.
HD61.H76 2009
658.15'5–dc22

2008052147

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

I dedicate this book to my entire support staff: my wife, Janet, and our children, Evan, Madeleine, and Steven.

Contents

Preface

I started writing this book in early 2008, well before the most serious period of the financial crisis. The original plan was to turn in my manuscript in December but, as the economic crisis developed, the publisher saw that a book about the failure of risk management might become more relevant to many readers. So, at my editor's urging, instead of writing a 50,000-word manuscript due by December, I wrote an 80,000-word manuscript by the end of October.

Although the financial crisis becomes an important backdrop for a book about risk management, I still wanted to write a much broader book than a reaction to the most recent disaster. This book should be just as relevant after the next big natural disaster, major product safety recall, or catastrophic industrial accident. Better yet, I hope readers see this book as a resource they need *before* those events occur. Risk management that simply reacts to yesterday's news is not risk management at all.

I addressed risk in my first book, *How to Measure Anything: Finding the Value of Intangibles in Business*. Risk struck me as one of those items that is consistently perceived as an *intangible* by management. In a way, they are right. A risk that something could occur—the probability of some future event—is not *tangible* in the same way as progress on a construction project or the output of a power plant. But it is every bit as measurable. Two entire chapters in the first book focused just on the measurement of uncertainty and risks.

Unfortunately, risk management based on actual measurements of risks is not the predominant approach in most industries. I see solutions for managing the risks of some very important problems that are in fact no better than astrology. And this is not a controversial position I'm taking. The flaws in these methods are widely known to the researchers who study them. The

message has simply not been communicated to the larger audience of managers.

In 1994, I developed a method I called *Applied Information Economics*, in part for the same reason that I wrote this and the previous book. I have watched consultants come up with a lot of half-baked schemes for assessing risks, measuring performance, and prioritizing portfolios with no apparent foundation in statistics or decision science. Arbitrary scoring schemes have virtually taken over some aspects of formalized decision-making processes in management. In other areas, some methods that do have a sound scientific and mathematical basis are consistently misunderstood and misapplied.

Of all the good, solid academic research and texts on risk analysis, risk management, and decision science, none seem to be directly addressing the problem of the apparently unchecked spread of pseudoscience in this field. In finance, Nassim Taleb's popular books, *Fooled by Randomness* and *The Black Swan*, have pointed out the existence of serious problems. But in those cases, there was not much practical advice for risk managers and very little information about assessing risks outside of finance. There is a need to point out these problems to a wide audience for a variety of different risks.

This book is somewhat more confrontational than my first one. No doubt, some proponents of widely used methods—some of which have been codified in international standards—might feel offended by some of the positions I am taking in this book. As such, I've taken care that each of the key claims I make about the weaknesses of some methods is supported by the thorough research of others, and not just my own opinion. The research is overwhelmingly conclusive—much of what has been done in risk management, when measured objectively, has added no value to the issue of managing risks. It may actually have made things worse.

Although the solution to better risk management is, for most, better quantitative analysis, a specialized mathematical text on the analysis and management of risks would not reach a wide enough audience. The numerous such texts already published haven't seemed to penetrate the management market, and I have no reason to believe that mine would fare any better. The approach I take here is to provide my readers with just enough technical information that they can make a 180-degree turn in risk management. They can stop using the equivalent of astrology in risk

management and at least start down the path of the better methods. For risk managers, mastering those methods will become part of a longer career and a study that goes beyond this book. This is more like a first book in astronomy for recovering astrologers—we have to debunk the old and introduce the new.

Douglas W. Hubbard

Acknowledgments

Many people helped me with this book in many ways. Some I have interviewed for this book, some have provided their own research (even some prior to publication), and others have spent time reviewing my manuscript and offering many suggestions for improvement. In particular, I would like to thank Dr. Sam Savage of Stanford University, who has been extraordinarily helpful on all these counts.

Daniel Kahneman
Tony Cox
Christopher ``Kip'' Bohn
Jim Deloach
Stephen Wolfram
Bob Clemen
Robin Dillon-Merrill
Karen Jenni
Rick Julien
David Budescu
Ray Covert
David Bearden
Jim Franklin
John Schuyler
Dennis William Cox
Jim Dyer
Steve Hoye
John Spangenberg
Jason Mewes
Dan Garrow
Reed Augliere
Fiona MacMillan
Bill Panning
Yook Seng Kong
Allen Kubitz
Andrew Braden
Rob Donat
Diana Del Bel Belluz
Andrew Freeman
Thompson Terry
Vic Fricas
Dr. Sam Savage
David Hubbard

PART ONE

An Introduction to the Crisis

CHAPTER 1

Healthy Skepticism for Risk Management

It is far better to grasp the universe as it really is than to persist in delusion, however satisfying and reassuring.

—Carl Sagan

Everything's fine today, that is our illusion.

—Voltaire

Any new and rapidly growing trend in management methods should be considered with healthy skepticism, especially when that method is meant to help direct and protect major investments and inform key public policy. It is time to apply this skepticism to the "risk management" methods meant to assess and then mitigate major risks of all sorts. Many of these methods are fairly new and are growing in popularity. Some are well-established and highly regarded. Some take a very soft, qualitative approach and others are rigorously quantitative. But for all of these methods, we have to ask the same, basic questions:

- Do any of these risk management methods work?
- Would anyone in the organization even know if they didn't work?
- If they didn't work, what would be the consequences?

For most organizations, the answers to these questions are all bad news. Natural, geopolitical, and financial disasters in the first few years of the 21st century have, perhaps only temporarily, created a new awareness of risk among the public, businesses, and lawmakers. This has spurred the development of several risk management methods, in both financial and non-financial sectors. Unfortunately, when these methods are measured rigorously, they don't appear to work. Most of the new non-financial methods are not based on any previous theories of risk analysis and there is no real, scientific evidence that they result in a measurable reduction in risk or improvement in decisions. Where scientific data does exist, the data shows that most methods fail to account for known sources of error in the analysis of risk or, worse yet, add error of their own. Even in the financial sector and other areas that use the most sophisticated, quantitative methods, there is a growing realization that certain types of systematic errors have undermined the validity of their analysis for years.

The answer to the second question (whether anyone would know that the risk management system has failed) is also *no*; most managers would not know what they need to look for to evaluate a risk management method and, more likely than not, can be fooled by a kind of "placebo effect"[1] and groupthink about the method. Even under the best circumstances, where the effectiveness of the risk management method itself was tracked closely and measured objectively, adequate evidence may not be available for some time. A more typical circumstance, however, is that the risk management method itself has no performance measures at all, even in the most diligent, metrics-oriented organizations. This widespread inability to make the sometimes-subtle differentiation between methods that work and methods that don't work means that ineffectual methods are likely to spread. Ineffectual methods may even be touted as "best practices" and, like a dangerous virus with a long incubation period, are passed from company to company with no early indicators of ill effects until it's too late.

Common Mode Failure

Finally, to answer the question about the consequences of unsound risk management methods, I'll use an example from a historic air-travel disaster to explain a concept called *common mode failure* (a concept from one of the more scientific approaches to risk analysis). In July 1989, I was the

commander of the Army Reserve unit in Sioux City, Iowa. It was the first day of our two-week annual training and I had already left for Fort McCoy, Wisconsin, with a small group of support staff (the "advance party"). The convoy of the rest of the unit was going to leave that afternoon, about five hours behind us. But just before the main body was ready to leave for annual training, the unit was deployed for a major local emergency.

United Airlines flight 232 to Philadelphia was being redirected to the small Sioux City airport because of serious mechanical difficulties. It crashed, killing 111 passengers and crew. Fortunately, the large number of emergency workers available and the heroic airmanship of the crew helped make it possible to save 185 onboard. Most of my unit spent the first day of our annual training collecting the dead from the tarmac and the nearby cornfields.

During the flight, the DC-10's tail-mounted engine failed catastrophically, causing the fast-spinning turbine blades to fly out like shrapnel in all directions. The debris from the turbine managed to cut the lines to *all three* redundant hydraulic systems, making the aircraft nearly uncontrollable. Although the crew was able to guide the aircraft in the direction of the airport by varying thrust to the two remaining wing-mounted engines, the lack of tail control made a normal landing impossible.

Aviation officials would refer to this as a "one-in-a-billion" event[2] and the media repeated this claim. But since mathematical misconceptions are common, if someone tells you that something that just occurred had merely a one-in-a-billion chance of occurrence, you should consider the possibility that they calculated the odds incorrectly.

The type of event that caused the crash is called a *common mode failure*, because a single event caused the failure of multiple components in a system. If they had failed independently of each other, the failure of all three would be extremely unlikely. But because all three hydraulic systems had lines near the tail engine, a single event could damage all of them. The common mode failure wiped out the benefits of redundancy.

Now consider that the cracks in the turbine blades would have been detected except for what the National Transportation Safety Board (NTSB) called "inadequate consideration given to human factors" in the turbine blade inspection process. Is human error more likely than one in a billion? Absolutely; in a way, that was an *even more common* common mode failure in the system.

But the common mode failure hierarchy could be taken even further. Suppose that the risk management method itself was fundamentally flawed. If that were the case, then perhaps problems in design and inspection procedures would be very hard to discover and much more likely to materialize. Now suppose that the risk management methods not just in one airline but in most organizations in most industries were flawed. The effects of disasters like Katrina and the financial crisis of 2008/9 could be inadequately planned for simply because the methods used to assess the risk were misguided. Ineffective risk management methods that somehow manage to become standard spread this vulnerability to everything they touch.

The ultimate common mode failure would be a failure of risk management itself. A weak risk management approach is effectively the biggest risk in the organization.

If the initial assessment of risk is not based on meaningful measures, the risk mitigation methods—even if they could have worked—are bound to address the wrong problems. If risk assessment is a failure, then the best case is that the risk management effort is simply a waste of time and money because decisions are ultimately unimproved. In the worst case, the erroneous conclusions lead the organization down a more dangerous path that it would probably not have otherwise taken.

The financial crisis occurring while I wrote this book was another example of a common mode failure that traces its way back to the failure of risk management of firms like AIG, Lehman Brothers, Bear Stearns, and the federal agencies appointed to oversee them. Previously loose credit practices and overly leveraged positions combined with an economic downturn to create a cascade of loan defaults, tightening credit among institutions, and further economic downturns. If that weren't bad enough, poor risk management methods are used in government and business to make decisions that not only guide risk decisions involving billions—or trillions—of dollars, but are also used to affect decisions that impact human health and safety.

What happened is history. But here are just a few more examples of major, risky decisions currently made with questionable risk assessment

methods, some of which we will discuss in more detail later. Any of these, and many more, could reveal themselves only after a major disaster in a business, government program, or even your personal life:

- The approval and prioritization of investments and project portfolios in major U.S. companies
- The evaluation of major security threats for business and government
- The decision to launch the space shuttle
- The approval of government programs worth many billions of dollars
- The determination of when additional maintenance is required for old bridges
- The evaluation of patient risks in health care
- The identification of supply chain risks due to pandemic viruses
- The decision to outsource pharmaceutical production to China

Clearly, getting any of these risks wrong would lead to major problems—as has already happened in some cases. The individual method used may have been sold as "formal and structured" and perhaps it was even claimed to be "proven." Surveys of organizations even show a significant percentage of managers who will say the risk management program was "successful" (more on this to come). Perhaps success was claimed for the reason that it helped to "build consensus," "communicate risks," or "change the culture."

Since the methods used did not actually measure these risks in a mathematically and scientifically sound manner, management doesn't even have the basis for determining whether a method works. Surveys about the adoption and success of risk management initiatives are almost always self-assessments by the surveyed organizations. They are not independent, objective measures of success in reducing risks. If the process doesn't correctly assess and mitigate risks, then what is the value of building consensus about it, communicating it, or changing the culture about it? Even if harmony were achieved, perhaps communicating and building consensus on the wrong solution will merely ensure that one makes the big mistakes faster and more efficiently.

Fortunately, the cost to fix the problem is almost always a fraction of a percent of the size of what is being risked. For example, a more realistic evaluation of risks in a large IT portfolio worth over a hundred million

dollars would not have to cost more than half a million—probably a lot less. Unfortunately, the adoption of a more rigorous and scientific management of risk is still not widespread. And for major risks such as those in the previous list, that is a big problem for corporate profits, the economy, public safety, national security, and you.

What Counts as Risk Management

There are numerous topics in the broad category of *risk management* but it is often used in a much narrower sense than it should be. When the term is used too narrowly, it is either because *risk* is used too narrowly, *management* is used too narrowly, or both.

If you start looking for definitions of *risk*, you will find many wordings that add up to the same thing, and a few versions that are fundamentally different. For now, I'll skirt some of the deeper philosophical issues about what it means (yes, there are some, but that will come later) and I'll avoid some of the definitions that seem to be unique to specialized uses. Chapter 5 is devoted to why the definition I am going to propose is preferable to various mutually-exclusive alternatives that each have proponents who assume their's is the "one true" definition.

For now, I'll focus on a definition that, although it contradicts some definitions, best represents the one used by well-established, mathematical treatments of the term (e.g. actuarial science), as well as any English dictionary or even how the lay-public uses the term (see the box below).

DEFINITION OF *RISK*

Long definition: The probability and magnitude of a loss, disaster, or other undesirable event

Shorter (equivalent) definition: Something bad could happen

The second definition is more to the point, but the first definition gives us an indication of how to quantify a risk. First, we can state a probability that

the undesirable event will occur. Also, we need to measure the magnitude of the loss from this event in terms of financial losses, lives lost, and so on.

The undesirable event could be just about anything, including natural disasters, a major product recall, the default of a major debtor, hackers releasing sensitive customer data, political instability around a foreign office, workplace accidents resulting in injuries, or a pandemic flu virus disrupting supply chains. It could also mean personal misfortunes, such as a car accident on the way to work, loss of a job, a heart attack, and so on. Almost anything that could go wrong is a risk.

Since risk *management* generally applies to a management process in an organization, I'll focus a bit less on personal risks. Of course, my chance of having a heart attack is an important personal risk to assess and I certainly try to manage that risk. But when I'm talking about the failure of risk management—as the title of this book indicates—I'm not really focusing on whether individuals couldn't do a better job of managing personal risks like losing weight to avoid heart attacks (certainly, most should). I'm talking about major organizations that have adopted what is ostensibly some sort of formal risk management approach that they use to make critical business and public policy decisions.

Now, let us discuss the second half of the phrase *risk management*. Again, as with *risk*, I find multiple, wordy definitions for *management*, but here is one that seems to represent and combine many good sources:

DEFINITION OF *MANAGEMENT*

Long definition: The planning, organization, coordination, control, and direction of resources toward defined objective(s)

Shorter, folksier definition: Using what you have to get what you need

There are a couple of qualifications that, while they should be extremely obvious, are worth mentioning when we put *risk* and *management* together. Of course, when an executive wants to manage risks, he or she actually

wishes to reduce it or at least not unduly increase it in pursuit of better opportunities. And since the current amount of risk and its sources are not immediately apparent, an important part of reducing or minimizing risks is figuring out where the risks are. Also, risk management must accept that risk is inherent in business and risk reduction is practical only up to a point. Like any other management program, risk management has to make effective use of limited resources. Putting all of that together, here is a definition (again, not too different in spirit from the myriad definitions found in other sources):

DEFINITION OF RISK MANAGEMENT

Long definition: The identification, assessment, and prioritization of risks followed by coordinated and economical application of resources to minimize, monitor, and control the probability and/or impact of unfortunate events

Shorter definition: Being smart about taking chances

Risk management methods come in many forms, but the ultimate goal is to minimize risk in some area of the firm relative to the opportunities being sought, given resource constraints. Some of the names of these efforts have become terms of art in virtually all of business. A popular, and laudable, trend is to put the word *enterprise* in front of *risk management* to indicate that it is a comprehensive approach to risk for the firm. *Enterprise risk management (ERM)* is one of the headings under which many of the trends in risk management appear. I'll call ERM a type of risk management *program*, because this is often the banner under which risk management is known. I will also distinguish programs from actual methods since ERM could be implemented with entirely different methods, either soft or quantitative.

The following are just a few examples of various management programs to manage different kinds of risks (*Note:* Some of these can be components of others and the same program can contain a variety of different methods):

- Enterprise risk management (ERM)
- Portfolio management or project portfolio management (PPM)
- Disaster recovery and business continuity planning (DR/BCP)
- Project risk management (PRM)
- Governance risk and compliance (GRC)
- Emergency/crisis management processes

Risk management includes analysis and mitigation of risks related to physical security, product liability, information security, various forms of insurance, investment volatility, regulatory compliance, actions of competitors, workplace safety, getting vendors or customers to share risks, political risks in foreign governments, business recovery from natural catastrophes, or any other uncertainty that could result in a significant loss.

Anecdote: The Risk of Outsourcing Drug Manufacturing

At a conference organized by the Consumer Health Products Association (a pharmaceutical industry association), I witnessed a chemical engineer describing a new risk management process he had developed for his firm. The risk analysis method was meant to assess an important and emerging risk in this field.

To control costs, this large pharmaceutical manufacturer was more frequently outsourcing certain batch processes to China. Virtually all of this manufacturer's competition was doing the same. But while the costs were significantly lower, they had a concern that batches from China might have additional quality control issues over and above those of batches manufactured here in the United States. These concerns were entirely justified.

The conference was in October 2007, and earlier that year there had already been several widely publicized product safety incidents with goods produced in China. In June, there was a toxin found in toothpaste and lead found in toys produced in China. Then there was tainted pet food that killed as many as 4,000 pets. There was even the disturbing case of "Aqua Dots," the children's craft-beads that stuck together to make different designs. The coating of these beads could metabolize in the stomach to produce gamma hydroxy butyrate—the chemical used in date-rape drugs.

Except for me, almost all of the audience were chemists, chemical engineers, and industrial engineers. They were previously listening to extremely technical sessions on sheer stress of particles in various processing equipment, yield curves, and mechanical details of drug packaging. There was no shortage of scientific thinkers and, from what I could tell, no timidity about mathematical models.

Yet, when the presenter was explaining the details of his company's new method for analyzing the risk of batches outsourced to China, I saw none of the hard science and skeptical peer-review that seemed common in the other sessions. He was describing a method based on a subjective "weighted score."[3] In it, several "risk indicators" were each scored on a scale of 1 to 5. For example, if the manufacturer already produces a similar, but not identical, drug, it might get a low risk score of 2 on the indicator called "proven technical proficiency." If it was inspected by and got a positive evaluation from the Chinese health agency, but was not yet inspected by the Food and Drug Administration, then it might get a 4 on the "formal inspections" indicator. If the components of the drug required certain special safety controls that would be harder to outsource, then it might score as a higher risk in other areas. Each of these scores was based on the judgments of a team assembled to make these evaluations.

Then these scores were each multiplied by a weight of somewhere between 0.1 and 1.0 and then all of the weighted scores were totaled. The total of the weighted score might be 17.5 for one outsourcing strategy, 21.2 for another, and so on. The team that chose the scores also chose the weights and, again, it was based only on subjective judgments. The team further separated the resulting scores into various stratifications of risk that would, apparently, have some bearing on the decision to use a particular China-based source for a drug. For example, risk scores of over 20 might mean "Extremely high risk: Find an alternative"; 10 to 19 might mean "High risk: Proceed only with increased quality assurance," and so on.

When the engineer had finished describing the approach, I noticed that several heads in the room turned to me expecting some response. Earlier that day, I had given the keynote address describing, among other things, how risk can be quantified in a mathematically and scientifically meaningful way. Perhaps some were implementing something similar in their firms and were curious to see whether I would endorse it, but I suspect it was more likely they were expecting a criticism.